MAINTAINING LARGE SOFTWARE REPOSITORIES WITH SLINK

Alva L. Couch
Associate Professor of
Electrical Engineering
and Computer Science
Tufts University
couch@eecs.tufts.edu

Greg Owen
Xerox Information Systems
gowen@xis.xerox.com

Our Repositories

- One architecture per repository.
- Software ready to run.
- Long repository lifetime (5+ years).
- Many maintainers and staff changes.
- Continual additions and deletions.
- Multiple co-resident versions.
- Multiple user environments.
- NFS-mounted or replicated.

FILESYSTEM ROT: GRADUALLY THE REPOSITORY IS CORRUPTED BY

- **Conflicting** root actions.
- **Frequent** modification.
- **Disappearing** maintainers.
- **Untrained** maintainers.
- **Inadequate** documentation.

OUR GOALS

- **Flexible** management tools.
- **Predictable** behavior.
- **Minimal** training period.
- **Fluid** administrative policy.
- **Voluntary** compliance.

Enforce administrative consistency by making it easier to be consistent than inconsistent!

RELATED WORK

- NIST Depot.
- CMU Depot.
- UMASS Depot-lite.

All assume that:

- Software is organized in **packages**.
- Packages are manipulated as **units**.

SLINK

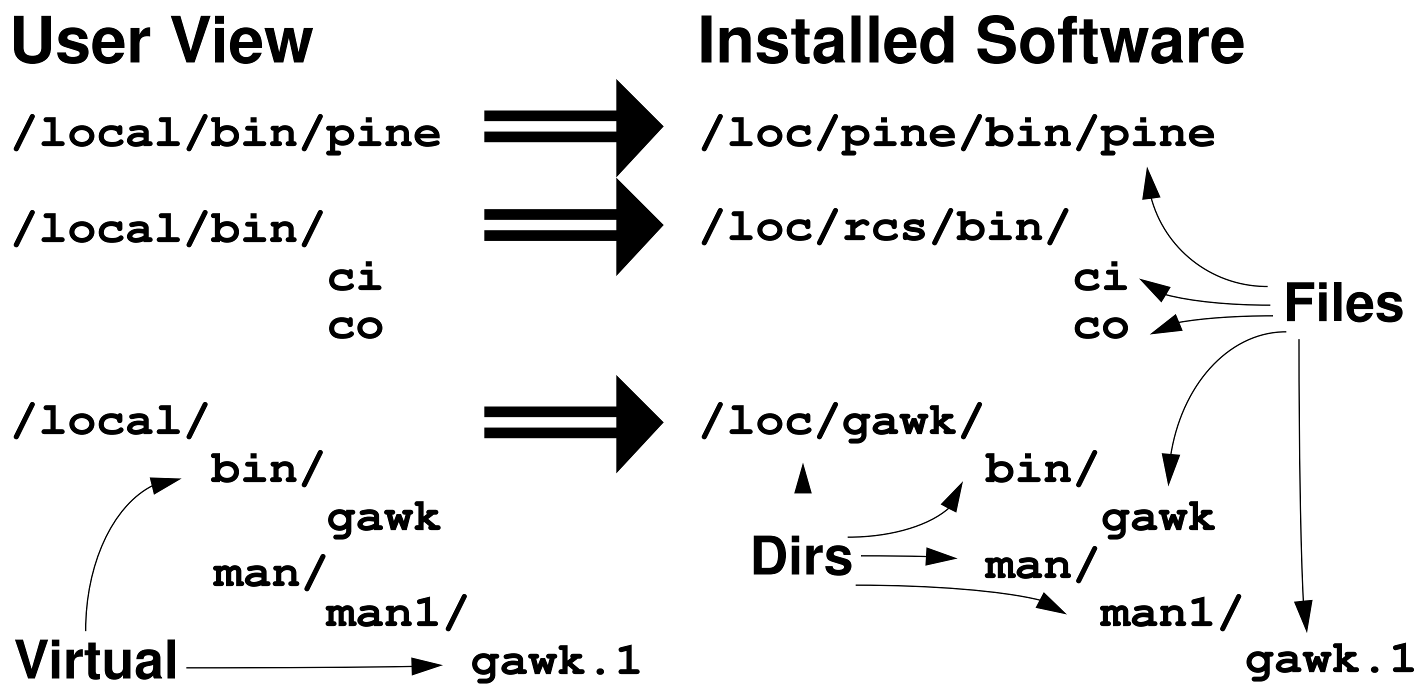
- General-purpose UNIX utility.
- Generalization of symbolic links.
- Forms unions of directory hierarchies.
- Table-driven for documentation.
- 1900 lines of PERL-4 (also PERL-5).

LINKS AND SLINKS

- A **slink A=>B** is an assertion that filesystem node **A** appears to contain the contents of node **B**.
- If **B** is a file, then **A=>B** is a simple symbolic link **A->B**.
- If **B** is a directory, then **A** can be a directory containing links to the contents of **B**, but perhaps more.

USING SLINKS

Slinks specify user views of software:



IMPLEMENTING SLINKS

- Slinks are formed **incrementally**.
- When a link must point to two directories, **brachiate** it into a directory of links to both.
- When a directory of links is equivalent with a single link, **condense** it into a single link.

Brachiation and Condensation

```
/local/                 /loc/gawk/
     bin/                       bin/
         gawk                       gawk
     man/                       man/
         man1/                      man1/
              gawk.1                     gawk.1
```

Brachiation **Condensation**

```
/local/                 /loc/gawk/
     bin/                       bin/
         gawk                       gawk
     man/                       man/
         man1/                      man1/
              gawk.1                     gawk.1
```

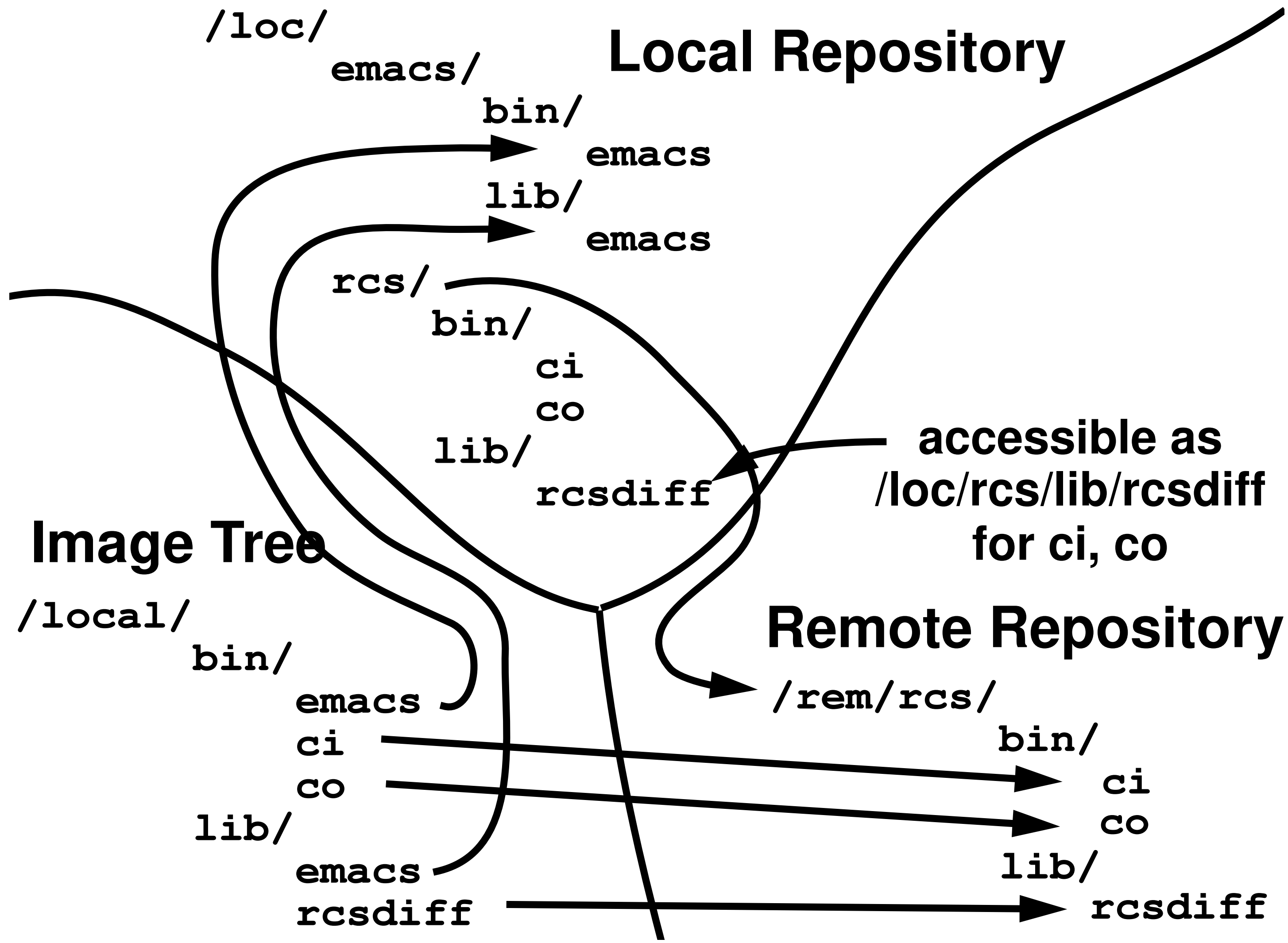
USES OF SLINK

- **Creating** user environments from package hierarchies.
- **Augmenting** vendor-supplied filesystems.
- **Constructing** tiered-memory repository hierarchies.

AUGMENTING VENDOR-SUPPLIED FILESYSTEMS

- **Problem:** 10 Alpha 3000's with too little local disk for all of **`/usr`**.
- **Solution:** on each machine:
- **Install** a minimal **`/usr`** that will fit.
- **Mount** a master **`/usr`** as **`/lusr`**.
- **Slink** **`/usr`** => **`/lusr`** to add the extra files to **`/usr`** (use **protect** to protect existing things in **`/usr`**).

TIERED REPOSITORY STRUCTURE



PROBLEMS IN SLINKING

- link traversal **overhead**:
  will using links take too much time?
- image tree **destruction**:
  will SLINK destroy important files?
- automount tree **confusion**:
  will SLINK create unmountable references?

AVOIDING LINK TRAVERSAL OVERHEAD

- SLINK's image tree contains only one symbolic link in any pathname.
- Even so, it is best to avoid mounting image trees via NFS and create one on each host.
- In practice, however, no one notices either symbolic or NFS lookup overhead.

PROTECTING IMAGES

- SLINK has four **protection modes**
  - **protect**: don’t touch existing links, directories, or files.
  - **relink**: change links or convert them to directories.
  - **redirect**: convert between links and directories.
  - **remove**: change anything.

AUTOMOUNT TREES

- Old automounters use symlinks into **`/tmp_mnt`** to simulate mounts.
- Left alone, SLINK would:
  - resolve names in automounted repositories into **`/tmp_mnt`**.
  - modify **`/tmp_mnt`** in changing an automounted image.
- Must specify prefixes to avoid in computing true pathnames.

SLINK DOES NOT:

- **enforce** package structure.
- **distribute** packages.
- **track** environment variables necessary for package use.

IN PRACTICE...

- Use of SLINK has spread 'like a disease'.
- It takes an hour or so to understand SLINK and policy and begin installing conforming applications.
- Filesystems remain rot-free even in the hands of beginners.
- Users like being able to find source trees by listing link contents.

CONCLUSIONS

- SLINK is a powerful tool for UNIX filesystem maintenance which allows:
- scalable administration
- fluid and rapid response.
- use and training of unskilled maintainers.
- rot-free repositories.

AVAILABILITY

The newest SLINK will soon be available at ftp.eecs.tufts.edu:/pub/slink-2.0.tar.Z. This version contains most of the improvements suggested in the paper.

Currently revision 1.2 is available and 2.0 is being tested at several beta sites.